Automatic stimulus evaluation depends on goal-relevance

Adriaan Spruyt[1], Helen Tibboel[2], Maarten De Schryver[1] & Jan De Houwer[1]

[1]Ghent University

[2]Erasmus University Rotterdam

Corresponding author:

Adriaan Spruyt

Department of Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent, Belgium

TEL: ++32-9-264-86-18

FAX: ++32-9-264-64-89

E-mail: Adriaan.Spruyt@UGent.be

Abstract

To examine whether automatic stimulus evaluation is dependent upon goal-relevance, participants were presented with a mixture of (a) goal-induction trials to create a set of goal-relevant and goal-irrelevant stimuli and (b) evaluative priming trials to capture the automatic evaluation of these stimuli as good or bad. In line with our predictions, a reliable evaluative priming effect was obtained only for stimuli that were relevant for the goal-induction task. Implications for the use of the evaluative priming paradigm as an assessment tool and the replicability of the evaluative priming effect in the absence of dimensional overlap between the prime set an the target set are discussed.

Automatic stimulus evaluation depends on goal-relevance

Automatic stimulus evaluation, defined as the fast, efficient, unintentional, uncontrollable, and/or unconscious evaluation of a stimulus as good or bad, plays a profound role in everyday life (Bargh, 2014). It influences the products we buy (e.g., Maison, Greenwald, & Bruin, 2004; Prestwich, Hurling, & Baker, 2011), the politicians we vote for (e.g., Arcuri, Castelli, Galdi, Zogmaister, & Amadori, 2008; Friese, Bluemke, & Wanke, 2007), the way we interact with people from other social groups (e.g., McConnell & Leibold, 2001; Rudman & Ashmore, 2007), and even the degree to which we engage in risky (e.g., Czopp, Monteith, Zimmerman, & Lynam, 2004; Traczyk & Zaleskiewicz, 2016) or addictive behavior (e.g., Marhe, Waters, van de Wetering, & Franken, 2013; Payne, Govorun, & Arbuckle, 2008).

One of the most well-known paradigms to capture automatic stimulus evaluation is the evaluative priming paradigm, introduced by Fazio, Sanbonmatsu, Powell, and Kardes (1986). In a typical evaluative priming study, participants judge the evaluative connotation of target stimuli that appear one by one in the center of a computer screen. Crucially, each of these targets is preceded by a brief presentation of a positive or negative prime stimulus. Results consistently show that participants are faster and more accurate to respond to targets that are preceded by primes with the same evaluative connotation as compared to targets that are preceded by primes with a different evaluative connotation (i.e., the so-called evaluative priming effect), even under strict automaticity conditions (for an overview see Herring et al., 2013; for an extensive treatment of the automaticity concept, see Moors, 2016; Moors & De Houwer, 2006).

The use of an evaluative categorization task is somewhat problematic, however, as it necessarily implies that participants are required to adopt an explicit evaluative processing goal. Such a requirement is typically missing in everyday life and, therefore, questions can be raised concerning the ecological validity of evaluative priming data obtained with the (standard) evaluative categorization task (Bargh, Chaiken, Raymond, & Hymes, 1996). To resolve this problem, a number of researchers have attempted to obtain the evaluative priming effect using a simple word-pronouncing task or picture-naming task. As an example, consider the evaluative priming study reported by Spruyt and Hermans (2008). They presented participants with four blocks of pronunciation trials in which each prime stimulus (i.e., real-life color pictures) and each target stimulus (i.e., words) was presented only once. The authors observed a clear-cut evaluative priming effect, suggesting that the evaluative priming effect can indeed arise in the absence of an evaluative processing mindset (e.g., Bargh et al., 1996).

The observations reported by Spruyt and Hermans (2008) are also important for a number of other reasons. First, although the overall evaluative priming effect did reach significance, planned comparisons revealed that the effect was reliable in the first block of trials only. The evaluative priming effect in the word-pronunciation task thus seems to dissipate rapidly across successive blocks, although it is still unclear exactly why this is the case. Second, given that (a) no stimulus was presented more than once within each block of trials and (b) a reliable effect was found in the first block, the findings of Spruyt and Hermans (2008) imply that the occurrence of the evaluative priming effect in the word-pronunciation task is not critically dependent upon a high degree of stimulus repetition (see Klauer & Musch, 2001). Finally, because participants are required to respond on the basis of the identity of the targets when pronouncing target words, the word-pronunciation task is characterized by an absence of dimensional overlap between the prime set and the response set (Kornblum, Hasbroucq, & Osman, 1990). As a result, the findings of Spruyt and Hermans (2008) imply that processes other than direct response facilitation and/or interference contribute to the emergence of the evaluative priming effect, an hypothesis that has been contested by several authors (e.g., De Houwer, Hermans, Rothermund, & Wentura, 2002; Klauer, Rossnagel, & Musch, 1997; Rothermund & Werner, 2014; Werner & Rothermund, 2013).

However, while several studies showing evaluative priming of pronunciation/naming responses appeared in the literature (e.g., Bargh et al., 1996; De Houwer, Hermans, & Spruyt, 2001; Everaert, Spruyt, & De Houwer, 2011; Schmitz & Wentura, 2012; Hermans, Smeesters, De Houwer, & Eelen, 2002; Spruyt, Hermans, De Houwer, & Eelen, 2002; Hermans, De Houwer, & Eelen, 1994, 2001; Spruyt, De Houwer, Everaert, & Hermans, 2012; Spruyt, De Houwer, & Hermans, 2009; Spruyt, Hermans, De Houwer, & Eelen, 2004; Spruyt, Hermans, De Houwer, Vandromme, & Eelen, 2007; Wentura & Frings, 2008; for related findings see Gast, Werner, Heitmann, Spruyt, & Rothermund, 2014; Spruyt, 2014; Spruyt & Tibboel, 2015), several authors reported that they were unable to replicate this effect (e.g., Becker, Klauer, & Spruyt, 2016; Klauer, Becker, & Spruyt, 2016; Klauer & Musch, 2001; Spruyt, Hermans, Pandelaere, De Houwer, & Eelen, 2004). Given the relatively large number of studies showing both reliable effects and null findings, it seems unlikely that the mixed pattern of findings simply resulted either from Type-I errors or Type-II errors only. Instead, it seems more likely that the occurrence of the evaluative priming effect in the pronunciation/naming task is dependent upon moderating factors.

One potential moderator is the degree to which evaluative relatedness between the primes and the targets is confounded with variations in associative relatedness, that is, the degree to which the activation of one concept (e.g., 'Laurel') will call to mind another concept (e.g., 'Hardy') due to their repeated temporal or spatial co-occurrence (Spruyt, Hermans, De Houwer, et al., 2004). It is a well-known fact that associative relations between a prime stimulus and a target stimulus can influence the speed of responding in a sequential priming paradigm, irrespective of whether there is dimensional overlap between the prime set and the response set (Hutchison, 2003; Neely, 1991). If it is assumed that associative relations are more likely to occur between stimuli that share a common evaluative meaning than between stimuli that are characterized by a conflicting evaluative meaning (Hermans et al., 2002; Klauer et al., 2016; Werner, von Ramin, Spruyt, & Rothermund, in press), it could be argued that uncontrolled, associative relatedness was the driving force behind the evaluative priming effects obtained with the pronunciation/naming task. In line with this reasoning, Klauer et al. (2016) failed to replicate the findings reported by Spruyt and Hermans (2008) when controlling for associative relatedness. Likewise, Werner et al. (in press) demonstrated that associative relatedness can produce normal or even reversed evaluative priming effects depending on the nature of the confound between evaluative relatedness and associative relatedness. In sum, there are good reasons to assume that associative relatedness is an important confounding factor that has been overlooked in evaluative priming research, not only in studies that relied on the use of the pronunciation/naming task but also in dozens of other studies in which the classic evaluative priming paradigm was used.

However, there is also evidence suggesting that an explanation in terms of uncontrolled associative relatedness will probably not suffice to explain the full pattern of results obtained with variants of the pronunciation/naming task. As an example, consider the findings reported by Hermans et al. (2002). They manipulated evaluative and associative relatedness orthogonally in a word-pronunciation task (i.e., Experiment 4) and obtained a reliable interaction between these two factors. Crucially, this interaction resulted from the fact that a reliable evaluative priming effect emerged only for prime-target pairs that were *unrelated* in terms of associative relatedness. No effect whatsoever emerged using prime-target pairs that were clearly related in terms of associative relatedness. Therefore, if anything, this study seems to provide positive evidence *against* the hypothesis that the evaluative priming effect in the word-pronunciation task can result only from (uncontrolled) variations in associative relatedness between the primes and the targets.

The available data thus seem to suggest that, in addition to associative relatedness, other moderating factors must have been at play in earlier studies showing reliable evaluative priming effects in the absence of overlap between the prime set and the response set. In this article, we argue that automatic stimulus evaluation in the absence of an explicit evaluative processing mindset may be dependent upon the degree to which a stimulus is relevant for a particular goal. This hypothesis was inspired by prior research showing that goals can increase the salience as well as the cognitive accessibility of information related to that goal (e.g., Kiefer, 2008; Anderson, Laurent, & Yantis, 2011a, 2011b; Levine & Edelstein, 2010; Moors, Spruyt, & De Houwer, 2010). In addition, it has been shown that goal-relevant information captures attention in an automatic fashion (Vogt, De Houwer, Moors, Van Damme, & Crombez, 2010), even when this information is presented together with threatening events (Vogt, De Houwer, Crombez, & Van Damme, 2013). Crucially, automatic stimulus evaluation is dependent upon attention assignment. The more attention is assigned to a given stimulus, the more likely that the evaluative tone of this stimulus is processed, even under automaticity conditions (De Houwer & Randell, 2002; Musch & Klauer, 2001; Simmons & Prentice, 2006). Accordingly, it could be hypothesized that automatic evaluative stimulus processing may be more pronounced for goal-relevant stimuli as compared to goal-irrelevant stimuli, especially if the explicit requirement to focus attention on the evaluative stimulus dimension is missing.

This hypothesis has the potential to shed new light on the mixed results found with the pronunciation/naming task. If generic stimulus materials are used, the degree to which these materials are personally relevant may fluctuate across studies depending on participant characteristics and/or the experimental setup. For example, first-year psychology students who participate in exchange for course credit may be motivated to learn about psychological research and might therefore have a personal interest in the experimental procedures and stimulus materials used. In contrast, if one uses stimulus materials linked to the (chronic) goals of the individual participants (e.g., pictures of spiders for a spider phobic), the evaluative priming effect in the pronunciation/naming task may be less volatile. The observation that one can use evaluative priming scores obtained with the naming task to predict relapse in abstinent smokers (Spruyt et al., 2015) is in line with this reasoning. After all, what is more goal-relevant for an abstinent smoker than a stimulus linked to the act of smoking a cigarette (for related findings, see Descheemaeker, Spruyt, & Hermans, 2014; Spruyt, Hermans, De Houwer, Vandekerckhove, & Eelen, 2007; Vandromme, Hermans, & Spruyt, 2011; but see Vanaelst, 2016)?

To test the hypothesis that the evaluative priming effect in the naming task is dependent upon goal-relevance, we adopted the task-switching approach developed by Vogt et al. (2013). Participants were thus asked to switch between goal-induction trials and evaluative priming trials that were presented in a random, intermixed order. The evaluative priming trials (71.43 %) were aimed at capturing automatic evaluative stimulus processing in the absence of (a) an explicit evaluative processing goal and (b) dimensional overlap between the prime set and the response set. More specifically, in line with Spruyt and Hermans (2008), participants were presented with four blocks of picture-word pronunciation trials and each target stimulus was presented exactly once within each block of trials. On the goal-induction trials (28.57 %), participants performed a go/no-go task in which prime stimuli were followed by the presentation of a target letter (i.e., the letter ‘E’ or ‘F’). Participants were asked to categorize the target letters as fast as possible, but only if the preceding prime stimulus was one of four specific prime stimuli. On all other letter-judgment trials (i.e., 12 in each block), participants were expected to withhold their response. Goal-relevant information can be defined as information that is an element of some plan sufficient to achieve a goal (Gorayska & Lindsay, 1993; Levine & Edelstein, 2010). Given that the go-cues were relatively rare (i.e., 4 out of 16 goal-induction trials per block), the most effective strategy to perform well on the go/no-go task (i.e., the goal) was to not respond unless a go cue was presented (i.e., the plan). In line with this plan, the instructions specified that it would be too difficult to remember each individual cue for the go/no-go task and, therefore, that the best strategy to perform well on this task was to focus attention on the subset of four primes that served as go-cues. In sum, the identity of the go-cues was goal-relevant because this information was an element of the plan to achieve a good performance in the go/no-go task.

Most importantly, the exact make-up the experimental design ensured that the go cues were goal-relevant throughout the entire experimental session. The four primes that served as go cues on the goal-induction trials (i.e., two positive and two negative, randomly selected at random for each individual participant) and another subset of four primes that served as no-go cues on the goal-induction trials (i.e., two positive and two negative, also selected at random for each individual participant) were used as primes in the actual evaluative priming trials. Because (a) the goal-induction trials and the evaluative priming trials were presented in a random order and (b) the nature of the required response (i.e., pronunciation vs. letter discrimination) was cued only after the presentation of the prime stimuli (i.e., by the presentation of a target word or a letter, respectively), we expected participants to discriminate between go cues and no-go cues throughout the entire experimental session. As a result, based on the assumption that automatic evaluative stimulus processing is dependent upon goal-relevance, we expected to observe a reliable (or more pronounced) evaluative priming effect for the go cues (i.e., the goal-relevant priming condition) as compared to the no-go cues (i.e., the goal-irrelevant priming condition). In addition, based on the observation that the evaluative priming dissipates rather rapidly across successive blocks of trials (i.e., Spruyt & Hermans, 2008), we expected this effect to be most pronounced in the first block of trials. In sum, the crucial statistical test for the present study is the three-way interaction between the factors block, goal-relevance, and evaluative congruence.

Importantly, the goal-relevant condition and the goal-irrelevant condition were equated in terms of associative relatedness. Moreover, we ensured that all prime-targets pairs identified by Klauer et al. (2016) as being related in a non-evaluative, associative manner were excluded a priori. Accordingly, should we observe an evaluative priming effect for goal-relevant primes only, it would be extremely hard to entertain the idea that such an effect is nothing else but a mere by-product of associative relations between the prime set and the target set.

In sum, the purpose of the present research is therefore two-fold. First, we sought to demonstrate that automatic evaluative stimulus processing is dependent upon goal-relevance. Second, we sought to demonstrate that a reliable evaluative priming effect can be obtained in the absence of dimensional overlap between the prime set and the response set even if non-evaluative, associative relatedness is properly controlled for.

Method

Participants.

Participants were 77 students at Ghent University. They all received course credit in exchange for their participation. All participants were Dutch speakers and had normal or corrected-to-normal vision. All participants gave informed consent prior to participation. Two participants did not complete the experimental procedure, for different reasons. One participant did not complete the experimental procedure because the computer program was unable to detect her verbal responses due to an experimenter error (i.e., use of an incorrect version of the computer program). A second participant did not complete the experimental session because she failed to understand the instructions. Three additional participants were excluded because they were excessive outliers in terms of the number of trials on which the voice key was not activated appropriately (i.e., 28.13 %, 39.38 %, and 61.88 %; 2.5-*SD* cutoff = 26.47 %, overall mean = 4.43 %)[1]. Finally, we excluded the data of one additional participant because she was an outlier in terms of the overall mean response time (i.e., 713 ms; 2.5-*SD* cutoff = 710 ms; overall mean 532 ms)[2]. In the final sample (N = 71), 14 participants were male and 57 were female. The mean age in this sample was 18.66 years. Overall, participants had little or no prior experience participating in psychological experiments. Twelve participants (n = 12) took part in a psychological study for the first time. Other participants reported to have participated in one (n = 21), two (n = 23), three (n = 11), or between four and ten (n = 4) other studies. The Ethics Committee of the Faculty of Psychology and Educational Sciences of Ghent University approved this research.

Materials.

All stimulus materials were identical to those used by Spruyt and Hermans (2008) and Klauer et al. (2016). More specifically, the prime set included 30 positive (e.g., a kitten) and 30 negative (e.g., a corpse) color pictures (resolution 512 x 384 pixels), some of which originated from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2001). Norm data collected by Spruyt et al. (2002) confirmed that the mean valence rating of the positive pictures was significantly larger than zero, on a scale ranging from -5 to + 5, M = 2.22, SE = 0.11, $t(29)$ = 19.96, p < .001. Likewise, the mean valence rating of the negative pictures was significantly smaller than zero, M = -2.77, SE = 0.19, $t(29)$ = -14.65, p < .001. The target set included 20 positive words (e.g., ‘LIEFDE’, the Dutch word for ‘love’) and 20 negative words (e.g., ‘HAAT’, the Dutch word for ‘hate’), presented in Arial uppercase letters (font size 36, RGB 255, 255, 255). Norm data collected by Hermans and De Houwer (1994) confirmed

[1] Inclusion or exclusion of these participants did not influence the outcome of the main analyses. Note, however, that our design was quite complex (16 conditions) relative to the total number of evaluative priming trials (i.e., 160 in total). High proportions of erroneous voice key triggers thus resulted in a limited number of observations in at least a number of cells. In fact, for one of these excluded participants, there was just one observation in one cell of the design, thereby making it impossible to include the data of this participant in a classic repeated measures ANOVA (see below). Therefore, despite the fact that the inclusion or exclusion of these participants had virtually no impact on our main findings, there can be little doubt that the exclusion of these participants was indeed a necessity.

[2] Inclusion or exclusion of this participant does not affect the results reported below.

that the mean valence rating of the positive words was significantly larger than four, on a scale ranging from 1 to 7, $M = 5.77$, $SE = 0.26$, $t(19) = 6.94$, $p < .001$. Likewise, the mean valence rating of the negative words was significantly smaller than four, $M = 1.71$, $SE = 0.08$, $t(19) = 21.15$, $p < .001$. Finally, a plus-sign (+) presented in Arial uppercase letters (font size 28, RGB 255, 255, 255) was used as a fixation stimulus. All stimuli were presented in the center of a 27-inch computer monitor with a black background (100 Hz, screen resolution 1024 x 768, RGB 0, 0, 0). An Affect 4.0 program (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010) controlled the presentation of the stimuli as well as the registration of the response latencies. Two AZERTY keyboards were used, one for the participants and one for the experimenter. An external voice key that was connected to the parallel port of the computer was used to measure response latencies.

Procedure.

Participants were tested individually in a darkened room. Participants completed four blocks of 56 trials each. Each block consisted of a mixture of 16 goal-induction trials and, identical to Spruyt and Hermans (2008), 40 evaluative priming trials, presented in a random, intermixed order. The inter-trial interval (ITI) always varied randomly between 500 ms and 1500 ms.

For each individual participant, the experimental computer program selected three sets of pictures (random sampling without replacement). First, two positive and two negative pictures were selected to serve as goal-relevant primes. Second, four additional pictures (two positive, two negative) were selected to be used as goal-irrelevant primes. Third, eight additional pictures (4 positive and 4 negative) were selected to serve as primes for (a subset of) the goal-induction trials only (hereafter referred to as induction stimuli).

For the evaluative priming trials, each potential target was presented exactly once within each block. In contrast, the primes used for the evaluative priming trials consisted of a subset of 8 pictures (i.e., the four goal-relevant primes and the four goal-irrelevant primes), each of which was presented exactly five times within each block. For each individual participant, primes and targets were combined in a semi-random order so that (a) all combinations of positive and negative primes and targets occurred equally often and (b) all prime-target pairs that were identified by Klauer et al. (2016) as being related in a non-evaluative, associative manner were excluded. Similar to earlier priming studies conducted by the first author, each trial started with the 500-ms presentation of the fixation stimulus, followed by a blank screen for 500 ms. Next, a prime stimulus was presented for 200 ms, followed by an inter-stimulus interval of 50 ms. Finally, a target was presented until a pronunciation response was detected by the voice key. By pressing one of three keys of one of the two computer keyboards, the experimenter coded whether the microphone was triggered accurately and whether the pronunciation response was correct. There was no error feedback for the evaluative priming trials.

For the goal-induction trials, each of the goal-relevant primes, the goal-irrelevant primes, and the induction stimuli was presented once within each block of trials (i.e., 16 trials in total). Each of these stimuli was followed by the presentation of the letter E or F (Arial uppercase letters, font size 36, RGB 255, 255, 255). Participants were asked to indicate as fast as possible whether the letter E or F was presented, but *only if* this letter was preceded by a goal-relevant prime stimulus (hereafter referred to as go trials). They were asked to press a left key (²) for the letter E and the right key (-) for the letter F. If the picture preceding the letter was a goal-irrelevant prime stimulus or an induction stimulus, no response was required (hereafter referred to as no-go trials). The temporal details of the goal-induction trials were identical to the sequence of events on the evaluative priming trials, with the exception that (a) the letters E and F were presented for a maximum duration of 2000 ms and (b) that error feedback was presented. If participants executed a response on a no-go trial, the Dutch translation of 'INCORRECT! You were not allowed to respond!' was presented. Likewise, the Dutch translation of 'INCORRECT! You should have responded!' was presented if participants failed to execute a response on a go trial. Finally, if participants committed an error on a go-trial, the Dutch translation of 'INCORRECT! Wrong response!' was presented. All error messages were presented for 5000 ms in the center of the computer screen in Arial uppercase letters (font size 28, RGB 255, 0, 0).

Preceding the actual evaluative priming task, participants completed 3 phases aimed at establishing experimental differences in the goal-relevant nature of the primes. In the first phase, participants were presented with a self-paced (random) sequence of the 4 goal-relevant primes. They were asked to watch these stimuli attentively because these stimuli would serve as go cues for the letter-judgement task during the actual experimental phase. In the second phase, participants were presented with a self-paced (random) sequence of the 4 goal-irrelevant primes. It was explained that, in total, 12 pictures would serve as no-go cues for the letter-judgement task and that it would be 'too difficult' to memorize each of these 12 stimuli. It was emphasized that focusing on the 4 pictures presented during the first phase (i.e., the goal-relevant primes) would be sufficient to perform well on the letter-judgement trials. Nevertheless, allegedly as a random subset of the pool of no-go cues, the 4 goal-irrelevant prime stimuli were presented. As a result, the goal-relevant and goal-irrelevant prime set were equated in terms of presentation frequency and/or novelty. Finally, in the third and last phase, participants practiced the letter-judgment task. In total, participants completed 16 goal-induction trials and each of the 16 prime stimuli (i.e., the 4 goal-relevant prime stimuli, 4 goal-irrelevant prime stimuli, and the 8 induction stimuli) was presented exactly once. Identical to the experimental blocks, the ITI varied randomly between 500 ms and 1500 ms.

At the very end of the experiment, participants completed an explicit evaluative rating phase in which each of the goal-relevant and goal-irrelevant prime stimuli was presented exactly once, again in a self-paced, random sequence. Participants were asked to evaluate each stimulus on a 21-pont rating scale ranging from -10 to + 10. Finally, for exploratory reasons, participants were asked to completed the Need to Evaluate Scale developed by Jarvis and Petty (1996).

Results

Data analysis.

The raw of this experiment are available at https://doi.org/10.6084/m9.figshare.4750831.v1. Trials on which a voice key error (2.82 %) or an incorrect response (0.36 %) was registered were excluded. To reduce the impact of outlying values, values smaller than 150 ms (0.06 %) or larger than

2000 ms (0.04 %) were discarded (see Ratcliff, 1993).[3] Finally, given (a) that participants were required to switch between a pronunciation task and a goal-induction task and (b) that the pronunciation task was not practiced prior to the start of the experimental phase of the experiment, the first pronunciation trial of each experimental session was excluded from the analyses (i.e., 0.63 %).[4]

To ensure that our findings can be generalized both across participants and stimuli (see Baayen, Davidson, & Bates, 2008; Judd, Westfall, & Kenny, in press), we adopted a crossed linear mixed effects approach using the R package 'lme4' (Bates, Mächler, Bolker, & Walker, 2014; R version 3.4.0). Following Wolsiefer, Westfall, and Judd (in press), participants, primes, and the targets were included in the model as random factors (see also Judd, Westfall, & Kenny, 2012). Goal relevance, evaluative congruence, and block were included as fixed, effect-coded factors and for each of these it was tested whether by-participant, by-prime, and/or by-target random slopes were needed. The reported *p*-values for the fixed effects are based on a Type-III ANOVA using a χ^2-distribution as implemented in the R package 'car' (Fox & Weisberg, 2011). For each block and for each level of goal-relevance, planned comparisons of congruent and incongruent trials were calculated using dummy coding. It may be noted that a classic repeated-measures ANOVA, complemented with a series of planned comparisons, produced similar results (see Footnote 2). All R-scripts used for the data analysis are publically at https://doi.org/10.6084/m9.figshare.5057743.v1. As requested by a reviewer, we report exact *p*-levels unless $p < .001$.

---

[3] In earlier papers by the first author, far-out values were typically defined as values that deviated more than 2.5 standard deviations from the mean of an individual participant in a particular cell of the design. This approach is suboptimal for the present dataset as the number of observations in each cell of the design was quite limited (i.e., 10 observations at best, that is, if no trials were lost due to voice key errors and/or incorrect responses). Given that cutoff values based on standard deviations become very liberal in such a scenario, we reasoned that it was best to adopt fixed cutoff values for the present data set. The choice for a lower cutoff value of 150 ms was based on Klauer et al. (2016), who also implemented this cutoff value. The choice for an upper cutoff value of 2000 ms was based on the original study by Spruyt and Hermans (2008) who implemented a response deadline of 2000 ms. It may be noted, however, that all critical results were replicated if the 2.5 *SD*-cutoff criterion was implemented, with or without the prior use of the fixed cutoff values. Also note that different outlier elimination methods did result in different participants being identified as outliers in terms of their overall response latency (reassuringly, without any consequences for the critical findings, except for the degrees of freedom).

[4] Unless noted otherwise, the inclusion or exclusion of the first pronunciation trial did not affect the pattern of results.

Evaluative priming data.

The three random effects (i.e., participants, primes, and targets) explained 34.02 % of the total variance. By-participant random slopes were needed for goal relevance and block. Because these random effects were highly correlated, the random factor block was recoded as a binary factor (i.e., first block vs. the remaining blocks). There was no evidence suggesting that by-primes or by-target random slopes were needed.

The main effect of goal-relevance was reliable, $\chi^2(1) = 11.24$, $p < .001$. Participants were slower to respond after the presentation of a goal-relevant prime, $M = 535$ ms, 95% CI [517 ms, 552 ms], than after the presentation of a goal-irrelevant prime, $M = 526$ ms, 95% CI [508 ms, 543 ms]. They also responded increasingly faster across successive blocks, $\chi^2(3) = 300.40$, $p < .001$. In the first block, the mean response latency was 594 ms, 95% CI [574 ms, 615 ms]. In the second block, the mean response latency was 522 ms, 95% CI [504 ms, 539 ms]. In the third block, the mean response latency was 509 ms, 95% CI [491 ms, 526 ms]. Finally, in the fourth block, the mean response latency was 501 ms, 95% CI [484 ms, 518 ms]. The main effect of evaluative congruence was unreliable, $\chi^2 < 1$.

More importantly, in line with our predictions, the model revealed a significant three-way interaction between goal-relevance, block, and evaluative congruence, $\chi^2(3) = 10.10$, $p = .018$.[5] Table 1 presents the estimated means and evaluative priming effects for each cell of the design. As expected, the evaluative priming effect reached significance only in the goal-relevant condition of the first block, i.e., 17 ms, $\chi^2(1) = 9.14$, $p = .003$. In all other conditions, there was no evidence for an evaluative priming effect, all $\chi^2 < 1.85$ and all $p > .17$. There was also no evidence suggesting that the effect of congruence was different for different participants, primes, or targets, all $\chi^2 < 1.72$ and all $p > .40$.[6]

[5] The three-way interaction between goal-relevance, block, and evaluative congruence just missed significance if the first pronunciation was not excluded as a practice trial $\chi^2(3) = 7.28$, $p = .063$. In all other aspects, the inclusion or exclusion of the first pronunciation trial had no effect. More specifically, the evaluative priming effect did reach significance, as predicted, in in the goal-relevant condition of the first block, i.e., 18 ms, $\chi^2(1) = 9.43$, $p = .002$. In all other conditions, there was no evidence for an evaluative priming effect, all $|\chi^2| < 1.62$, all $p > .20$.
[6] The results of a classic repeated-measures ANOVA were virtually identical. Using the same outlier-elimination method as for the crossed linear mixed effects analysis, the three-way interaction between goal-relevance, block, and evaluative congruence was reliable, $F(2.41, 168.68) = 3.62$, $p = .010$ (Greenhouse-Geisser corrected), $\eta_p^2 = .05$. In line with the crossed linear mixed effects analysis, planned comparisons revealed that the evaluative

Explicit rating data.

The explicit valence ratings of the experimental primes were also analyzed using a crossed linear mixed effects approach. Participants and stimuli were included in the model as random factors. Goal relevance and stimulus valence were included as fixed, effect-coded factors. The results showed a strong effect of stimulus valence, $\chi^2(1) = 758.34$, $p < .001$. The mean rating of the positive primes was 5.70, 95% CI [5.06, 6.34]. The mean rating of the negative primes was -6.84, 95% CI [-7.48, -6.21]. This main effect of stimulus valence did not interact with the factor goal relevance, $\chi^2 < 1$. The main effect of goal-relevance was also unreliable, $\chi^2 < 1$.

Discussion

As expected, the goal-relevant primes produced a reliable evaluative priming effect in the first block of trials. No effects were obtained with the goal-irrelevant primes. This observation is important for a number of reasons. First, we are the first to demonstrate that automatic stimulus evaluation as captured by the evaluative priming paradigm is dependent on goal-relevance. Earlier research already showed that neutral stimuli can function as positive or negative prime stimuli if they indicate success or failure, respectively, on a goal-inducing task (Moors & De Houwer, 2001, 2005; Moors, De Houwer, & Eelen, 2004). The present research adds to this line of research by showing that goal-relevance can influence not only the evaluative meaning of neutral stimuli but also the degree to which participants engage in automatic evaluative processing of stimuli with an existing evaluative meaning.[7] Of course,

---

priming effect was reliable only in the goal-relevant condition of the first block (i.e., 17 ms), $t(70) = 2.01$, $p = .048$, $d = .24$ (all other $|t| < 1.25$, all other $p > .20$).

[7] In contrast to the studies performed by Moors and colleagues (Moors & De Houwer, 2001, 2005; Moors et al., 2004), the present experiment included both goal-relevant and goal-irrelevant primes. It was therefore possible to examine whether goal-relevance in itself is evaluated in a positive or negative manner. Accordingly, we conducted an exploratory analysis in which goal-relevant and goal-irrelevant primes were coded as positive and negative stimuli, respectively. Evaluative congruence and block were then included as fixed factors in a crossed linear mixed effects model whereas participants, primes, and the targets were again included as random factors. Preliminary results suggested that by-participant random slopes were needed for block (again coded as binary factor, see above) as well as by-target random slopes for congruence. In the final model, the interaction between congruence and block reached significance, $\chi^2(3) = 17.17$, $p < .001$. Follow-up analyses revealed that the effect of evaluative congruence was reliable in the first block (i.e., 14 ms), $\chi^2(1) = 11.77$, $p < .001$, suggesting that goal-relevant stimuli, irrespective of their intrinsic meaning, are evaluated in a positive manner. In the second block of trials, however, a reversed evaluative priming effect was obtained (i.e., -8 ms), $\chi^2(1) = 4.07$, $p = .043$. No effects were obtained in the third and fourth block, both $\chi^2 < 1$. The present data are therefore inconclusive concerning the evaluative meaning of goal-relevance stimuli.

the question now arises as to what mechanism is responsible for the observation that automatic stimulus evaluation is more pronounced for goal-relevant stimuli as compared to goal-irrelevant stimuli. As already noted above, there is evidence showing that goal-relevant stimuli attract (spatial) attention in an automatic fashion (e.g., Tibboel, Liefooghe, & De Houwer, 2016; Vogt et al., 2013; Vogt et al., 2010), potentially because they are salient (Anderson et al., 2011a, 2011b), arousing (e.g., Montagrin & Sander, 2016), and/or highly accessible in working memory (e.g., Maxcey-Richard & Hollingworth, 2013). Given that automatic evaluative stimulus processing increases as a function of (spatial) attention assignment (De Houwer & Randell, 2002; Musch & Klauer, 2001; Simmons & Prentice, 2006), one can argue that the impact of goal-relevance upon automatic evaluative stimulus processes is mediated by selective attention assignment. As an alternative explanation, however, one might argue that the intermixed presentation of go/no-go trials and evaluative priming trials resulted in a general readiness to respond to the goal-relevant primes (i.e., the go signals; see Schuch & Koch, 2003). Such a mechanism can also account for the present findings if one assumes that go signals are more likely to elicit responses in general, that is, including evaluative responses as captured by the evaluative priming paradigm. More research is thus in order to pinpoint the exact mechanism(s) that contributed to our findings. It may be noted, however, that we can already argue against the hypothesis that goal-relevance simply increases the extremity of a stimulus, as the evaluative ratings collected at the very end of the experiment showed no impact of the goal-relevance manipulation whatsoever.

Second, the present experiment replicates the observation by Spruyt and Hermans (2008) that the occurrence of the evaluative priming effect in in word-pronunciation task is restricted to a first block of trials (but see Klauer et al., 2016). One way to explain this finding relates to the fact that, both in the present study and the study reported by the Spruyt and Hermans (2008), each target stimulus was presented exactly once within each block. Target repetition was thus absent in the first block but increased as the experiment progressed. We suspect that target repetition results in an increase of the efficiency of the target identification process, thereby reducing the extent to which this process can be influenced by the evaluative match between the (goal-relevant) primes and the targets (for related arguments, see Malley & Strayer, 1995). Interestingly, this interpretation also resonates with earlier studies showing reliable evaluative priming effects in the word-pronunciation task if the swift and efficient identification of the target stimuli is hampered by a degradation manipulation (De Houwer et al., 2001). Further research would be required, however, to substantiate this interpretation.

Third, the present findings are difficult to explain in terms of associative relatedness. Both Klauer et al. (2016) and Werner et al. (in press) argued that evaluative and associative relatedness may have been confounded in earlier studies showing reliable evaluative priming effects in the absence of dimensional overlap between the prime set and the response set (see also Hermans et al., 2002). It is certainly true that associative relatedness is an important confounding factor that has been overlooked in evaluative priming research in general (but see Hermans et al., 2002). Still, the present findings demonstrate that an explanation in terms of uncontrolled, associative relatedness will probably fail to account for the full pattern of results reported in the literature. In the present study, we replicated Klauer et al. (2016) in the sense that we never presented prime-target pairs that were characterized by a high degree of associative relatedness. In addition, given the use of a within-participants design, the goal-relevant condition and the goal-irrelevant condition were equated perfectly in terms of non-evaluative, associative relatedness. Nevertheless, as predicted, a sizable evaluative priming effect was found in the first block of trials of the goal-relevant condition only. Accordingly, we can safely reject the hypothesis that *if* an evaluative priming effect is found in the absence of dimensional overlap, it must be due to uncontrolled associative relatedness.

Fourth, the present findings are also important for the discussion concerning the mechanisms that translate the outcome of the prime-evaluation process into an observable evaluative priming effect (Spruyt, Gast, & Moors, 2011). There is little doubt that direct response facilitation and/or interference plays a critical role in the emergence of the evaluative priming effect when participants are asked to evaluate the targets (e.g., De Houwer et al., 2002; Klauer et al., 1997; Klinger, Burton, & Pitts, 2000; Rothermund & Werner, 2014; Spruyt, Hermans, De Houwer, Vandromme, et al., 2007;

Werner & Rothermund, 2013). For the present data, however, such a mechanism can be ruled out as each target stimulus required a unique (pronunciation) response. Our findings thus corroborate earlier studies showing evaluative priming effects in the absence of dimensional overlap (e.g., Bargh et al., 1996; De Houwer et al., 2001; Everaert et al., 2011; Gast et al., 2014; Hermans et al., 1994, 2001; Schmitz & Wentura, 2012; Spruyt, 2014; Spruyt et al., 2012; Spruyt et al., 2009; Spruyt, De Houwer, Hermans, & Eelen, 2007; Spruyt & Hermans, 2008; Spruyt et al., 2002; Spruyt, Hermans, De Houwer, et al., 2004; Spruyt, Hermans, De Houwer, Vandromme, et al., 2007; Spruyt & Tibboel, 2015; Wentura & Frings, 2008). Taken together, this line of research strongly suggests that processes other than direct response facilitation and/or interference can contribute to the emergence of the evaluative priming effect. As an example of such an alternative explanation, it has been suggested that the perceptual encoding of a target stimulus is more cognitively demanding (and therefore more time-consuming) if it is preceded by an incongruent prime stimulus (Gast et al., 2014), at least under conditions that promote selective attention for the evaluative stimulus dimension. Further research would be needed, however, to firmly substantiate this idea.

Fifth, the present data underscore the importance of automatic stimulus evaluation for everyday life. Unlike the evaluative categorization task, stimulus valence is a task-irrelevant stimulus feature in the word-pronunciation task. The fact that an evaluative priming effect can be obtained under these conditions thus confirms the classic hypothesis that automatic stimulus evaluation is not contingent upon the presence of an explicit evaluative processing goal (e.g., Bargh et al., 1996; Hermans et al., 1994), even though the effect seems to dissipate quickly over time (i.e., the effect was reliable only in the first block of trials). At first sight, however, this conclusion may seem incompatible with the observation that the evaluative priming effect typically fails to replicate in the absence of dimensional overlap between the prime set and the target set unless the experimental conditions encourage participants to assign selective attention to the evaluative stimulus dimension (e.g., Everaert et al., 2011; Gast et al., 2014; Spruyt et al., 2009; Spruyt & Tibboel, 2015). To reconcile these findings, we argue that selective attention for the evaluative stimulus dimension may be a prerequisite for obtaining the evaluative priming effect only if the primed stimuli under investigation have no particular relevance for the participants. That is, as soon as a stimulus becomes goal-relevant, stimulus evaluation may take place in a truly unconditional, automatic fashion. One may argue, however, that the automaticity conditions realized in the present experiment were somewhat limited as participants were required to processes the primes intentionally. While this approach is certainly an important difference between the present study and more 'typical' pronunciation/naming studies, it must emphasized that the requirement to process the primes was limited to the *identity* of the primes. That is, similar to other naming/pronunciation studies, the evaluative tone of the primes was completely task-irrelevant. Moreover, the requirement to process the identity of the primes concerned both the goal-relevant and the goal-irrelevant primes. The observation that the occurrence of the evaluative priming effect was restricted to the goal-relevant condition thus implies that the requirement to process of the identity of the primes intentionally is in itself not sufficient to obtain a reliable evaluative priming effect. What does seem to matter is the degree to which a prime stimulus is relevant for achieving a particular goal. Still, in future studies, it might be interesting to manipulate the goal-relevance of the prime stimuli in a manner that does not require participants to process the identity of the primes intentionally.

Finally, the present findings have important implications for research concerning the predictive validity of implicit measures of evaluation (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). Spruyt et al. (2015) observed that evaluative priming scores obtained with the picture-picture naming task can be used to predict relapse in abstinent smokers. Given that smoke-related pictures must be goal-relevant for an abstinent smoker, this observation is perfect accordance with the idea that the evaluative priming effect in a pronunciation/naming task depends on goal-relevance (see also Descheemaeker et al., 2014; Spruyt, Hermans, De Houwer, Vandekerckhove, et al., 2007; Vandromme et al., 2011). However, as evidenced by a large-scale meta-analysis by Vanaelst (2016), research on the predictive validity of implicit measures of evaluation relies almost exclusively on healthy, student populations for whom the stimuli under investigation bear little personal relevance. As a result, so the present findings suggest, the field is at risk of reaching the wrong conclusions concerning the (applied) value of different (implicit) measures of evaluation. In conclusion, then, there is a pressing need for large-scale, applied studies in which stimuli are used that are truly personally relevant.

Author Note

Correspondence concerning this article should be addressed to Adriaan Spruyt, Department of Psychology, Ghent University, H. Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail may be sent to Adriaan.Spruyt@UGent.be. Preparation of this paper was supported by Methusalem Grant BOF16/MET_V/002 of Ghent University.

References

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *Plos One, 6*, ARTN e27926. DOI 10.1371/journal.pone.0027926

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 10367-10371. DOI 10.1073/pnas.1104047108

Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology, 29*, 369-387. DOI 10.1111/j.1467-9221.2008.00635.x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. DOI 10.1016/j.jml.2007.12.005

Bargh, J. A. (2014). Our unconscious mind. *Scientific American, 310*, 30-37

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104-128. DOI 10.1006/jesp.1996.0005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*

Becker, M., Klauer, K. C., & Spruyt, A. (2016). Is attention enough? A Re-examination of the impact of feature-specific attention allocation on semantic priming effects in the pronunciation task. *Attention, Perception, and Psychophysics, 78*, 396-402. DOI 10.3758/s13414-015-1046-7

Czopp, A. M., Monteith, M. J., Zimmerman, R. S., & Lynam, D. R. (2004). Implicit attitudes as potential protection from risky sex: Predicting condom use with the IAT. *Basic and Applied Social Psychology, 26*, 227-236. DOI 10.1207/s15324834basp2602&3_9

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition & Emotion, 16*, 643-666. Doi 10.1080/02699930143000419

De Houwer, J., Hermans, D., & Spruyt, A. (2001). Affective priming of pronunciation responses: Effects of target degradation. *Journal of Experimental Social Psychology, 37*, 85-91. DOI 10.1006/jesp.2000.1437

De Houwer, J., & Randell, T. (2002). Attention to primes modulates affective priming of pronunciation responses. *Experimental Psychology, 49*, 163-170. DOI 10.1027//1618-3169.49.3.163

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*, 347-368. DOI 10.1037/A0014211

Del Re, A. C. (2013). Compute.es: Compute Effect Sizes. R package version 0.2-2. URL: http://cran.r-project.org/web/packages/compute.es.

Descheemaeker, M., Spruyt, A., & Hermans, D. (2014). On the relationship between the indirectly measured attitude towards beer and beer consumption: The role of attitude accessibility. *Plos One, 9*, e95302. DOI 10.1371/journal.pone.0095302

Everaert, T., Spruyt, A., & De Houwer, J. (2011). On the (un)conditionality of automatic attitude activation: The valence proportion effect. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 65*, 125-132. DOI 10.1037/A0022316

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238. DOI 10.1037//0022-3514.50.2.229

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd ed.): Thousand Oaks, CA: Sage.

Friese, M., Bluemke, M., & Wanke, M. (2007). Predicting voting behavior with implicit attitude measures - The 2002 German parliamentary election. *Experimental Psychology, 54*, 247-255. DOI 10.1027/1618-3169.54.4.247

Gast, A., Werner, B., Heitmann, C., Spruyt, A., & Rothermund, K. (2014). Evaluative Stimulus (In)Congruency Impacts Performance in an Unrelated Task Evidence for a Resource-Based Account of Evaluative Priming. *Experimental Psychology, 61*, 187-195. DOI 10.1027/1618-3169/A000238

Gorayska, B., & Lindsay, R. (1993). The roots of relevance. *Journal of Pragmatics, 19*, 301-323. DOI 10.1016/0378-2166(93)90091-3

Hermans, D., & De Houwer, J. (1994). Affective and subjective familiarity ratings of 740 Dutch words. *Psychologica Belgica, 34*, 115-139

Hermans, D., De Houwer, J., & Eelen, P. (1994). The affective priming effect: automatic activation of evaluative information in memory. *Cognition & Emotion, 8*, 515-533. DOI 10.1080/02699939408408957

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion, 15*, 143-165. DOI 10.1080/0269993004200033

Hermans, D., Smeesters, D., De Houwer, J., & Eelen, P. (2002). Affective priming for associatively unrelated primes and targets. *Psychologica Belgica, 42*, 191-212

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., . . . Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin, 139*, 1062-1089. DOI 10.1037/a0031309

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin and Review, 10*, 785-813. DOI 10.3758/Bf03196544

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*, 172-194. DOI 10.1037/0022-3514.70.1.172

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem. *Journal of Personality and Social Psychology, 103*, 54-69. DOI 10.1037/a0028347

Judd, C. M., Westfall, J., & Kenny, D. A. (in press). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*. DOI 10.1146/annurev-psych-122414-033702

Kiefer, M. (2008). Constraining unconscious cognition: Top-down modulation of unconscious priming. *International Journal of Psychology, 43*, 727-727

Klauer, K. C., Becker, M., & Spruyt, A. (2016). Evaluative priming in the pronunciation task: A preregistered replication and extension. *Experimental Psychology, 63*, 70-78. DOI 10.1027/1618-3169/a000286

Klauer, K. C., & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 54*, 727-751. DOI 10.1080/02724980042000444

Klauer, K. C., Rossnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning Memory and Cognition, 23*, 246-255. DOI 10.1037/0278-7393.23.1.246

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*, 441-455. DOI 10.1037//0278-7393.26.2.441

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility - A model and taxonomy. *Psychological Review, 97*, 253-270. DOI 10.1037/0033-295x.97.2.253

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2001). International affective picture system (IAPS): Instruction manual and affective ratings. Gainsville, Florida: Center for Research in Physiology, University of Florida.

Levine, L. J., & Edelstein, R. S. (2010). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion, 23*, 833-875. DOI 10.1080/02699930902738863

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the implicit association test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology, 14*, 405-415. DOI 10.1207/s15327663jcp1404_9

Malley, G. B., & Strayer, D. L. (1995). Effect of stimulus repetition on positive and negative identity priming. *Perception and Psychophysics, 57*, 657-667. DOI 10.3758/Bf03213271

Marhe, R., Waters, A. J., van de Wetering, B. J. M., & Franken, I. H. A. (2013). Implicit and explicit drug-related cognitions during detoxification treatment are associated with drug relapse: An ecological momentary assessment study. *Journal of Consulting and Clinical Psychology, 81*, 1-12. DOI 10.1037/A0030754

Maxcey-Richard, A. M., & Hollingworth, A. (2013). The strategic retention of task-relevant objects in visual working memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 39*, 760-772. DOI 10.1037/a0029496

McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435-442. DOI 10.1006/jesp.2000.1470

Montagrin, A., & Sander, D. (2016). Emotional memory: From affective relevance to arousal. *Behavioral and Brain Sciences, 39*. DOI 10.1017/S0140525X15001879

Moors, A. (2016). Automaticity: componential, causal, and mechanistic explanations. *Annual Review of Psychology, Vol 67, 67*, 263-287. DOI 10.1146/annurev-psych-122414-033550

Moors, A., & De Houwer, J. (2001). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition & Emotion, 15*, 749-766

Moors, A., & De Houwer, J. (2005). Unintentional processing of motivational valence. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 58*, 1043-1063. DOI 10.1080/02724980443000467

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*, 297-326. DOI 10.1037/0033-2909.132.2.297

Moors, A., De Houwer, J., & Eelen, P. (2004). Automatic stimulus-goal comparisons: Support from motivational affective priming studies. *Cognition & Emotion, 18*, 29-54. DOI 10.1080/02699930244000462

Moors, A., Spruyt, A., & De Houwer, J. (2010). In search of a measure that qualifies as implicit: Recommendations based on a decompositional view of automaticity. In B. Gawronski & K. B. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and application*. New York: Guilford Press.

Musch, J., & Klauer, K. C. (2001). Locational uncertainty moderates affective congruency effects in the evaluative decision task. *Cognition & Emotion, 15*, 167-188

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Payne, B. K., Govorun, O., & Arbuckle, N. L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition and Emotion, 22*, 238-271. DOI 10.1080/02699930701357394

Prestwich, A., Hurling, R., & Baker, S. (2011). Implicit shopping: Attitudinal determinants of the purchasing of healthy and unhealthy foods. *Psychology & Health, 26*, 875-885. DOI 10.1080/08870446.2010.509797

Ratcliff, R. (1993). Methods for dealing with reaction-time outliers. *Psychological Bulletin, 114*, 510-532. DOI 10.1037/0033-2909.114.3.510

Rothermund, K., & Werner, B. (2014). Affective priming in the valent/neutral categorisation task is due to affective matching, not encoding facilitation: Reply to Spruyt. *Cognition and Emotion, 28*, 570-576. DOI 10.1080/02699931.2013.856778

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit association test. *Group Processes and Intergroup Relations, 10*, 359-372. DOI 10.1177/1368430207078696

Schmitz, M., & Wentura, D. (2012). Evaluative Priming of Naming and Semantic Categorization Responses Revisited: A Mutual Facilitation Explanation. *Journal of Experimental Psychology-Learning Memory and Cognition, 38*, 984-1000. DOI 10.1037/A0026779

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 92-105. DOI 10.1037/0096-1523.29.1.92

Simmons, J. P., & Prentice, D. A. (2006). Pay attention! Attention to the primes increases attitude assessment accuracy. *Journal of Experimental Social Psychology, 42*, 784-791. DOI 10.1016/j.jesp.2005.12.001

Spruyt, A. (2014). Attention please: Evaluative priming effects in a valent/non-valent categorisation task (reply to Werner & Rothermund, 2013). *Cognition and Emotion, 28*, 560-569. DOI 10.1080/02699931.2013.832653

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0 a free software package for implementing psychological and psychophysiological experiments. *Experimental Psychology, 57*, 36-45. DOI 10.1027/1618-3169/A000005

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition, 122*, 91-95. DOI 10.1016/j.cognition.2011.08.017

Spruyt, A., De Houwer, J., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language, 61*, 37-54. DOI 10.1016/j.jml.2009.03.004

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54*, 44-53. DOI 10.1027/1618-3169.54.1.44

Spruyt, A., Gast, A., & Moors, A. (2011). The sequential priming paradigm: A primer. In K. C. Klauer, C. Stahl & A. Voss (Eds.), *Cognitive methods in social psychology* (pp. 48-77). New York: Guilford.

Spruyt, A., & Hermans, D. (2008). Affective priming of naming responses does not depend on stimulus repetition. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 62*, 237-241. DOI 10.1037/1196-1961.62.4.237

Spruyt, A., Hermans, D., De Houwer, J., & Eelen, P. (2002). On the nature of the affective priming effect: Affective priming of naming responses. *Social Cognition, 20*, 227-256. DOI 10.1521/soco.20.3.227.21106

Spruyt, A., Hermans, D., De Houwer, J., & Eelen, P. (2004). Automatic non-associative semantic priming: Episodic affective priming of naming responses. *Acta Psychologica, 116*, 39-54. DOI 10.1016/j.actpsy.2003.12.012

Spruyt, A., Hermans, D., De Houwer, J., Vandekerckhove, J., & Eelen, P. (2007). On the predictive validity of indirect attitude measures: Prediction of consumer choice behavior on the basis of affective priming in the picture-picture naming task. *Journal of Experimental Social Psychology, 43*, 599-610. DOI 10.1016/j.jesp.2006.06.009

Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., & Eelen, P. (2007). On the nature of the affective priming effect: Effects of stimulus onset asynchrony and congruency proportion in naming and evaluative categorization. *Memory & Cognition, 35*, 95-106. DOI 10.3758/Bf03195946

Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., & Eelen, P. (2004). On the replicability of the affective priming effect in the pronunciation task. *Experimental Psychology, 51*, 109-115. DOI 10.1027/1618-3169.51.2.109

Spruyt, A., Lemaigre, V., Salhi, B., Van Gucht, D., Tibboel, H., Van Bockstaele, B., . . . Nackaerts, K. (2015). Implicit attitudes towards smoking predict long-term relapse in abstinent smokers. *Psychopharmacology, 232*, 2551-2561. DOI 10.1007/s00213-015-3893-2

Spruyt, A., & Tibboel, H. (2015). On the automaticity of the evaluative priming effect in the valent/non-valent categorization task. *Public Library of Science ONE*, e0121564. DOI 10.1016/j.jbtep.2015.04.003

Tibboel, H., Liefooghe, B., & De Houwer, J. (2016). Attention to future actions: the influence of instructed S-R versus S-S mappings on attentional control. *Psychological Research, 80*, 905-911. DOI 10.1007/s00426-015-0695-4

Traczyk, J., & Zaleskiewicz, T. (2016). Implicit attitudes toward risk: the construction and validation of the measurement method. *Journal of Risk Research, 19*, 632-644. DOI 10.1080/13669877.2014.1003957

Vanaelst, J. (2016). *Automatic evaluative stimulus processing as a function of feature-specific attention allocation: New directions in measuring and changing implicit evaluations.* (Unpublished doctoral dissertation), Ghent University, Ghent, Belgium.

Vandromme, H., Hermans, D., & Spruyt, A. (2011). Indirectly Measured Self-esteem Predicts Gaze Avoidance. *Self and Identity, 10*, 32-43. DOI 10.1080/15298860903512149

Vogt, J., De Houwer, J., Crombez, G., & Van Damme, S. (2013). Competing for attentional priority: temporary goals versus threats. *Emotion, 13*, 587-598. DOI 10.1037/a0027204

Vogt, J., De Houwer, J., Moors, A., Van Damme, S., & Crombez, G. (2010). The automatic orienting of attention to goal-relevant stimuli. *Acta Psychologica, 134*, 61-69. DOI 10.1016/j.actpsy.2009.12.006

Wentura, D., & Frings, C. (2008). Response-bound primes diminish affective priming in the naming task. *Cognition and Emotion, 22*, 374-384. DOI 10.1080/02699930701446064

Werner, B., & Rothermund, K. (2013). Attention please: No affective priming effects in a valent/neutral-categorisation task. *Cognition and Emotion, 27*, 119-132. DOI 10.1080/02699931.2012.711744

Werner, B., von Ramin, E., Spruyt, A., & Rothermund, K. (in press). Does sunshine prime loyal… or summer? Effects of associative relatedness on the evaluative priming effect in the valent/neutral categorization task. *Cognition and Emotion*

Wolsiefer, K., Westfall, J., & Judd, C. M. (in press). Modeling stimulus variation in three common implicit attitude tasks. *Behavior Research Methods*. DOI DOI 10.3758/s13428-016-0779-0

Table 1. Estimated mean response latencies (in ms), evaluative priming effects (in ms), effect sizes, and confidence intervals as a function of goal-relevance, block, and evaluative congruence.

	Block			
Evaluative congruence	Block 1	Block 2	Block 3	Block 4
Goal-relevant primes				
Congruent	597	526	509	504
Incongruent	613	531	508	500
Evaluative priming effect	17	5	- 1	- 3
$\chi^2(1)$	9.14	1.08	< 1	< 1
p	.003	.299	.912	.517
Effect size *r*	.36	.12	.01	.08
Lower bound 95% CI	.13	-.12	-.22	-.16
Upper bound 95% CI	.55	.35	.25	.31
Goal-irrelevant primes				
Congruent	587	516	511	496
Incongruent	580	513	506	503
Evaluative priming effect	- 7	- 3	- 5	7
$\chi^2(1)$	1.46	< 1	< 1	1.84
p	.227	. 526	.352	.174
Effect size r	.14	.08	.11	.16
Lower bound 95% CI	-.10	-.16	-.13	-.08
Upper bound 95% CI	.37	.31	.34	.38

Note: Effect sizes and corresponding confidence intervals computed on the basis of the reported χ^2 statistics using the R package ‘compute.es’ (Del Re, 2013).